Available online at www.sciencedirect.com

SCIENCE DIRECT

Developmental Biology 283 (2005) 474 – 485

DEVELOPMENTAL BIOLOGY

www.elsevier.com/locate/ydbio

Genomes & Developmental Control

Identification of a conserved 125 base-pair *Hb9* enhancer that specifies gene expression to spinal motor neurons

Takahiro Nakano[a], Martha Windrem[b], Vincenzo Zappavigna[c], Steven A. Goldman[a,b,*]

[a]*Department of Neurology and Neuroscience, Cornell University Medical Center, New York, NY 10021, USA*
[b]*Department of Neurology, University of Rochester Medical Center, Rochester, NY 14642, USA*
[c]*Department of Animal Biology, University of Modena and Reggio Emilia, 41100 Modena, Italy*

Received for publication 19 January 2005, revised 8 April 2005, accepted 12 April 2005
Available online 23 May 2005

Abstract

The homeobox gene *Hb9* is expressed selectively by motor neurons (MNs) in the developing CNS. Previous studies have identified a 9-kb 5′ fragment of the mouse *Hb9* gene that is sufficient to direct gene expression to spinal MNs in vivo. Here, we sought to identify more discrete MN-specifying elements, using homology searches between genomic sequences of evolutionarily distant species. Based on homology screening of the mouse and human *Hb9* promoters, we identified a 3.6-kb *Hb9* enhancer that proved sufficient to drive MN-specific *lacZ* expression. We then compared mouse, human, and pufferfish (*Fugu rubripes*) genomic sequences, and identified a conserved 438-bp sequence, consisting of noncontiguous 313-bp and 125-bp fragments, residing within the 3.6-kb *Hb9* enhancer. The zebrafish (*Danio rerio*) *Hb9* genomic region was then found to have two identical copies of the 125-bp sequence, but no counterpart for the 313-bp sequence. Transgenic analysis showed that the 125-bp alone was both necessary and sufficient to direct spinal MN-specific *lacZ* expression, whereas the 313-bp sequence had no such enhancer activity. Moreover, the 125-bp *Hb9* enhancer was found to harbor two *Hox*/*Pbx* consensus-binding sequences, mutations of which completely disrupted thoracolumbar *Hb9* expression. These data suggest that *Hox*/*Pbx* plays a critical role in the segmental specification of spinal MNs. Together, these results indicate that the molecular pathways regulating *Hb9* expression are evolutionarily conserved, and that MN-specific gene expression may be directed and achieved using a small 125-bp 5′ enhancer.

Keywords: Motor neurons; Promoters; Spinal muscular atrophy; Amyotrophic lateral sclerosis

Introduction

In the vertebrate CNS, the specification of neural identity is initiated by humoral inductive factors, that impose a specific profile of transcription factor expression on neural progenitor cells, thereby restricting their phenotypic differentiation (Goridis and Brunet, 1999; Jessell and Melton, 1992). In the developing spinal neuroepithelium, motor neuron (MN) progenitors arise in part in response to the ventralizing action of Sonic hedgehog (Shh) (Briscoe and Ericson, 2001). The specification of MN progenitors by Shh is mediated through the patterned expression of homeodomain (HD) and basic helix–loop–helix (bHLH) transcription factors; these function primarily as transcriptional repressors (Muhr et al., 2001), whose cross-regulatory interactions establish distinct progenitor domains (Briscoe et al., 1999, 2000; Jessell, 2000; Vallstedt et al., 2001). Through this general scheme, MN progenitors are restricted to a narrow region of the ventral neural tube termed the pMN domain (Briscoe et al., 2000; Jessell, 2000; Pierani et al., 2001). Within this domain, two HD proteins, *Nkx6.1* and *Pax6*, and a bHLH protein, *Olig2*, characterize MN

* Corresponding author. Department of Neurology, University of Rochester Medical Center, 601 Elmwood Avenue/MRBX, Box 645, Rochester, NY 14642, USA. Fax: +1 585 506 1957.
E-mail address: Steven_Goldman@urmc.rochester.edu (S.A. Goldman).

0012-1606/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ydbio.2005.04.017

progenitors. Together, these proteins serve to initiate the expression of distinct MN transcription factors, which include the HD protein HB9.

HB9 is expressed selectively by post-mitotic spinal MNs in the developing vertebrate CNS, and serves as a marker for the MN phenotype (Arber et al., 1999; Tanabe et al., 1998). Genetic studies in mice have suggested its importance in the consolidation and maintenance of MN identity (Arber et al., 1999; Thaler et al., 1999). A 5′ 9-kb *Hb9* promoter has been shown to drive MN-specific expression in vivo (Arber et al., 1999; Wichterle et al., 2002). Nonetheless, the regulatory control of *Hb9* gene expression is only poorly understood. In this study, we sought to identify *cis*-acting regulatory elements of the *Hb9* gene specifically active in MNs, so as to predict factors that might regulate MN induction. Using cross-species homology analysis with enhancer screening by transgenesis, we identified a highly conserved 125-bp *cis*-acting regulatory sequence, which appears to direct gene expression to spinal MNs. Moreover, by site-directed mutagenesis, we found that disrupting *Hox*/*Pbx* binding sites within this 125-bp *Hb9* enhancer completely abolished *β-gal* expression in the thoracic and lumber spinal cord, without affecting the reporter gene expression in the cervical levels of the cord. Thus, the 5′ 9-kb *Hb9* promoter harbors a highly conserved 125-bp element that directs vertebrate MN gene expression, and *Hox*/*Pbx* binding sites within this element appear necessary for thoracolumbar MN specification.

Materials and methods

Homology screening

A 9-kb 5′ non-coding fragment of the mouse *Hb9* gene (a kind gift from Dr. Thomas Jessell) was sequenced. The corresponding human genome sequence was obtained from Ensembl genome browser (www.ensemble.org). Blast 2 sequences program (www.ncbi.nih.gov/blast/bl2seq/bl2.html) was used to identify conserved non-coding regions of the *Hb9* gene between mouse and human. Ensembl genome browser was also used to identify evolutionarily conserved non-coding sequences of the *Hb9* genes among mouse, rat, human, pufferfish, and zebrafish, and then multiple sequence alignments were conducted using MacVector (Accelrys).

Transgenic constructs

Construct #1 (*3.6-kb Hb9 enhancer-β-globin-lacZ*) was generated by inserting a 3.6 kb *Not*I/*Sse*8387I fragment derived from a 9-kb 5′ non-coding fragment of the mouse *Hb9* gene into the *Not*I/*Pst*I sites of the reporter construct BGZA, which contains the *β-globin* minimal promoter, *lacZ* gene, and *SV40* polyadenylation cassette (a gift of Dr. Jane Johnson). Construct #2 (*5.4-kb Hb9 promoter-lacZ*) contains a 5.4-kb *Sse*8387I/*Pme*I fragment of the 9-kb *Hb9* promoter region into the *Pst*I/*Sma*I sites of *placZpA*. Construct #3 (*438-bp Hb9 enhancer-β-globin-lacZ*) was obtained by cloning 313-bp and 125-bp fragment derived from Constructs #4 and #5, respectively, to BGZA.

Construct #4 (*313-bp Hb9 enhancer-β-globin-lacZ*): A 313-bp fragment was PCR amplified with the primers:

5′-ATAGCATAGCGGCCGCTGAATAAATTTAAGCAGGCT-3′, 5′-GCTCTAGAAGCCCCATCCCCCTTCAAT-3′,

and cloned into BGZA.

Construct #5 (*125-bp Hb9 enhancer-β-globin-lacZ*): A 125-bp fragment was PCR amplified with the primers:

5′-GACTAGTAGAGTGGTTAGCTGATGAAT-3′, 5′-TCACCCGGGTCTAATCAGCCTGCCTAGCT-3′,

and cloned into BGZA. Constructs #4 and #5 contain three copies of the 313-bp and 125-bp fragments, respectively, in order to drive the reporter gene expression more efficiently.

Construct #6: The site-directed mutagenesis construct was generated according to Ho's method (Ho et al., 1989) using the primers:

5′-ATAGCATAGCGGCCGCTGAATAATTTAAGCAGGCT-3′,
5′-T**CG**TT**C**GTTTTTGTCAA**CG**CA**CG**AGCTAACCACTCTGGCTGGA-3′,
5′-T**CG**TG**CG**-TTGACAAAAAC**G**AA**CG**AGCTTCGAGCTTTATTGGGAAACAGGT-3′,
5′-TCACCCGGGTCTAATCAGCCTGCCTAGCT-3′.

All PCR-amplified fragments were verified by sequencing.

Production and genotyping of transgenic mice

Transgenes for injection were separated from vector sequences using 1% Seakim LE agarose (BMA), purified on QIAquick gel extraction kit (Qiagen), precipitated in injection buffer. Transgenic mice were generated by standard procedures (Hogan et al., 1994) using fertilized eggs from B6D2F1 9C57BL/6 × DBA) crosses. Transgenics were identified by PCR with the *lacZ* primers 5′-CGAGTGTGATCATCTGGTCG-3′ and 5′-TTACCTTGTGGAGCGACATC-3′ using genomic DNA extracted from yolk sacs or tails.

lacZ detection

β-gal staining of whole-mount embryos

Staged embryos were dissected from the uterus in cold PBS and fixed for 45 min at 4°C in fixing solution (1%

formaldehyde, 0.2% glutaraldehyde, 0.02% NP-40 in PBS pH 7.4). Whole-mount *β-gal* staining was performed as described (Ovitt et al., 1997).

β-gal staining of spinal cord tissue

To examine *β-gal* staining in adult spinal cord, stable transgenic lines carrying Construct #1 (*3.6-kb Hb9 enhancer-β-globin-lacZ*) were established. Genotyped adult transgenic mice were killed and perfusion fixed (1% formaldehyde, 0.2% glutaraldehyde, 0.02% NP-40 in PBS pH 7.4), and their spinal cords were removed. After *β-gal* staining (Ovitt et al., 1997), post-fixation was performed in the same fixative for 1–2 h at 4°C.

Immunohistochemistry

Embryos or adult spinal cords were cryosectioned at 15 μm, permeabilized with PBS, 0.1% saponin and 1% NGS, and blocked with PBS, 0.05% saponin and 5% NGS, each for 30 min. Sections were labeled with mouse anti-Islet1 (39.4D5, 1:100, DSHB), mouse anti-MNR2/Hb9 (5C10, 1:3, DSHB), or anti-ChAT (1:1000, Chemicon) overnight at 4°C. Species- and isotype-specific biotin-labeled secondary antibodies were applied at 1:100 for 2 h at room temperature. Staining was performed using the ABC kit (Vector) and DAB (Sigma).

Protein production and electophoretic mobility shift assay (EMSA)

PBX1 and HOX proteins used in this study were produced in vitro from the corresponding pSG5-derived expression vectors using a T7 polymerase-based coupled transcription–translation reticulocyte lysate system (Promega, Madison, WI). EMSA were performed according to the instruction of the DIG gel shift kit (Roche). Briefly, 2 μl of double-stranded DIG-labeled oligonucleotides was incubated with 2 μl of in-vitro-translated proteins in 20 μl of binding buffer [20 mM HEPES pH 7.6, 1 mM EDTA, 10 mM $(NH_4)_2SO_4$, 1 mM DTT, 0.2% (w/v) Tween 20, 30 mM KCl in the presence of 1 μg of poly (dI–dC) and 0.1 μg of poly L-lysine) for 30 min at room temperature. Pre-binding reactions of unlabeled competitor oligonucleotides at ~100-fold molar excess occurred for 15 min at room temperature. The incubation mixture was resolved by electrophoresis and served for chemiluminescent detection according to recommended conditions of the DIG gel shift kit (Roche). DIG-labeled DNA probes used were generated by annealing complementary oligonucleotides:

Wild-type oligonucleotide probe (40 bp)
5′-GTACGTTAGCTGATGAATTGACAAAAACTAATCAGCTTTA-3′
Mutant oligonucleotide probe (40 bp)
5′-GTACGTTAGCTCGTGCGTTGACAAAAACGAACGAGCTTTA-3′

Results

Identification of a 3.6-kb MN enhancer within the Hb9 locus

Previous work identified a 9-kb 5′ fragment of the mouse *Hb9* gene which is sufficient to direct spinal MN expression in vivo (Arber et al., 1999; Wichterle et al., 2002). To define the *cis*-acting regulatory sequences that govern the expression of *Hb9* by MNs, we first sought to identify conserved regions between mouse and human *Hb9* genomic sequences. Such conservation across species of non-coding sequences in the vicinity of a gene frequently marks the presence of regulatory elements (Muller et al., 2002). We thus compared the genomic sequences of mouse and human 9 kb *Hb9* promoter (see Materials and methods). Five clusters of remarkably high nucleotide sequence conservation (81% to 94%) were found (Fig. 1A; conserved regions <70 bp not shown). Three of these were located within a 5′ 3.6-kb fragment, whereas the other two were within the 5.4-kb *Hb9* promoter region.

To identify which of these elements directed *Hb9* expression, we established two transgenes, each of which was placed in control of *lacZ* (Fig. 1B). Transgene #1, containing the 3.6-kb *Hb9* fragment, was tested for enhancer activity by placing it 5′ of the *β-globin* basal promoter (P/*β-globin*) in a *lacZ* reporter vector, BGZA. This vector contains *lacZ* placed under the control of human P/*β-globin*, with an *SV40* polyadenylation cassette; in transgenics, it is inactive in the absence of an enhancer (Helms et al., 2000; Yee and Rigby, 1993). We designated the resultant plasmid E/*Hb9* (3.6 kb):*lacZ*. For transgene #2, the 5.4-kb *Hb9* promoter was used intact, directly driving *lacZ* with a polyA cassette, yielding the plasmid P/Hb9 (5.4 kb):*lacZ*.

Using these constructs, we examined gene expression patterns in transgenic embryos at 10.0–11.0 dpc. P/*Hb9* (5.4 kb):*lacZ* (transgene #2), did not activate *lacZ* expression in the neural tube (Fig. 1D). In contrast, E/*Hb9* (3.6 kb):*lacZ* (transgene #1) exhibited activity similar to that of transgenic embryos with the entire 9-kb *Hb9* promoter at E10.5–11.0 (Figs. 1C, 2A–C) (Arber et al., 1999). These results were consistent in at least 4 embryos bearing each transgene. In the E/*Hb9* (3.6 kb) mice, *lacZ* expression was noted in the ventral spinal cord at virtually all cervical, thoracic, and lumbar levels. Transverse sections revealed *lacZ* in the ventral horns and axons extending from the areas (Fig. 2D), suggesting the motor neuron specificity of E/*Hb9* (3.6 kb) at E10.5–11.0. These $lacZ^+$ cells co-expressed the MN transcription factors, *Islet-1* (Fig. 2E) and Hb9 (Fig. 2F). Together, these results indicated that, within the embryonic spinal cord, the 3.6-kb promoter fragment of *Hb9* specifically drives transgene expression in motor neurons. At E11.5, *β-gal* was noted within the trigeminal and facial motor nuclei as well as olfactory epithelium (Fig. 2C), where endogenous HB9

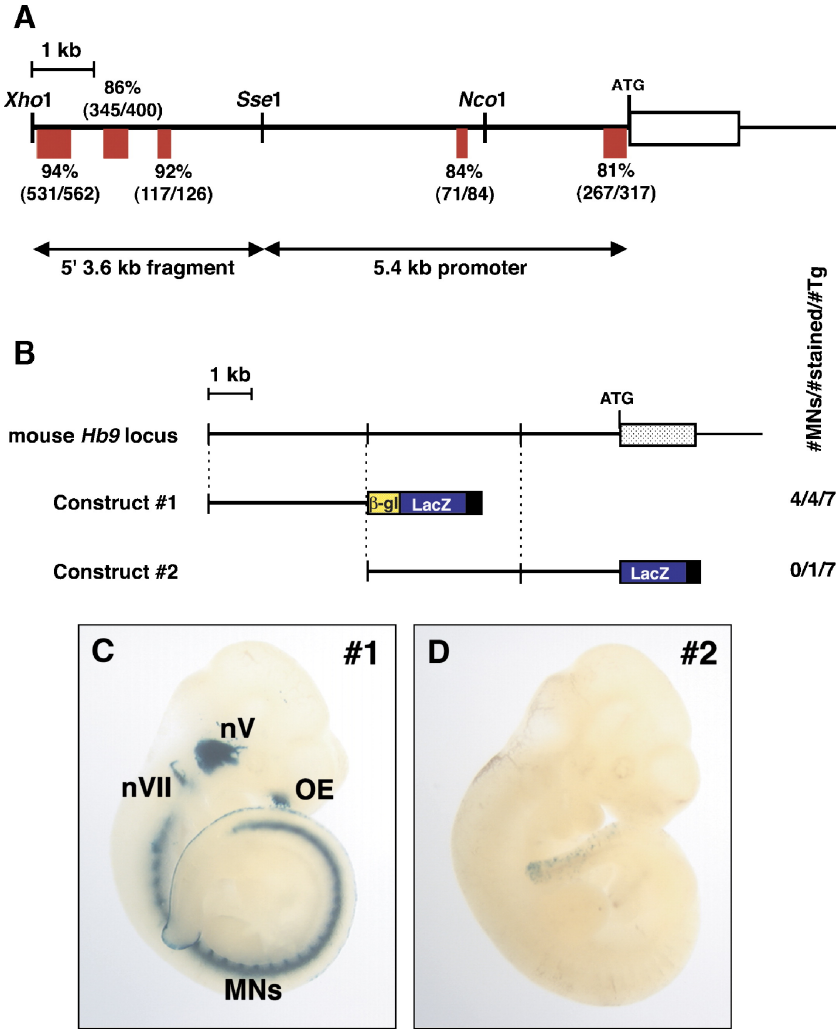


Fig. 1. A 5′ 3.6-kb fragment within the 9 kb *Hb9* promoter regulates *Hb9* gene expression. (A) Diagram of the *Hb9* locus representing the *Hb9* coding region (open box) and highly conserved sequences between the human and mouse *Hb9* loci (red boxes). The number and percentage of conserved nucleotides are shown. Scale bar = 1 kb. The position of the two fragments (3.6 kb and 5.4 kb) used to analyze gene expression activity in the *Hb9* region is indicated. (B) Transgenic constructs used to study gene expression activity. Statistical overview of transgenic embryos is indicated. The transgenic embryos carrying construct #1, but not #2, showed virtually the same spinal *β-gal* expression pattern with ones carrying MN-specific 9 kb *Hb9* promoter. *β-gl*, *β-globin* basal promoter in the BGZA reporter construct. (C, D) Representative pictures of E10.5 whole-mount stained embryos containing construct #1 (C) and #2 (D), respectively. OE, olfactory epithelium; nV, trigeminal nerve; nVII, facial nerve; MNs, spinal motor neurons.

protein is not expressed (Thaler et al., 1999), indicating some degree of ectopic expression by the 3.6-kb *Hb9* enhancer at later time points.

The establishment of stable transgenic lines of E/*Hb9* (3.6 kb):*lacZ* mice allowed us to examine gene expression in the adult spinal cord as well. Several litters from each of two independent transgenic lines were analyzed 8–12 weeks after birth. As in embryos, E/*Hb9*-driven *lacZ* was restricted to magnocellular neurons that reside in the median and lateral motor columns of the ventral spinal cord (Figs. 3A, C). $lacZ^{+}$ cells extended axons from ventral horns (Fig. 3B) and co-expressed ChAT (Fig. 3D), suggesting their MN phenotype. No differences were seen between the two lines. Thus, the 3.6-kb *Hb9* enhancer continues to direct MN-specific gene expression in the adult spinal cord.

The 3.6-kb Hb9 enhancer harbors evolutionarily conserved 313-bp and 125-bp fragments

To identify *cis*-acting elements that might directly regulate MN specification, we next used cross-species homology screening to identify conserved elements within the 3.6-kb *Hb9* enhancer. Comparison of mouse, rat, human, pufferfish (*Fugu rubripes*) 3.6-kb genomic sequences revealed two blocks of high homology, of 313 bp (region A, red arrow) and 125 bp (region B, blue arrowhead) (Fig. 4A). We found that both regions A and B were located within the 5′ regulatory region of the *Hb9* genes, confirmed by perfect conservation of the homeobox regions in these species (Fig. 4B). Overall, regions A and B exhibited 74% and 82% homology between mouse and *Fugu*, respectively (Figs. 4C–E). In *Fugu*, these two conserved regions were located within the first 2 kb

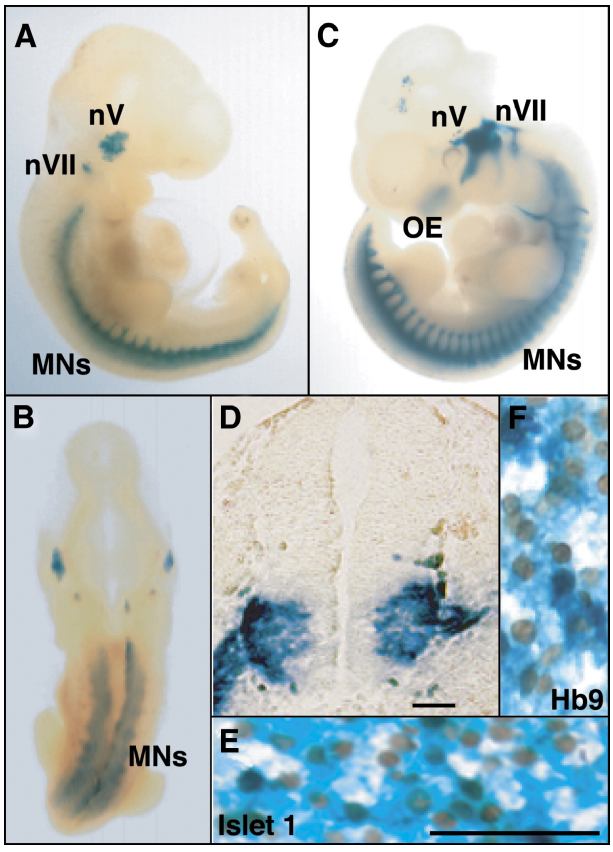


Fig. 2. Transgene expression in whole-mount mouse embryos. (A–C) *β-galactosidase* (*β-gal*) staining in E10 (A, B) and E11(C) E/*Hb9* (3.6 kb):*lacZ* transgenic embryos. Lateral (A, C) and dorsal (B) views. (D) A cross-section of E10.5 thoracic spinal cord showed *β-gal* expression in the ventrolateral cord and spinal roots. (E–F) Coronal sections of E/*Hb9* (3.6 kb):*lacZ* E10.5 spinal cord stained for Islet-1 (E) and Hb9 (F) proteins. OE, olfactory epithelium; nV, trigeminal nerve; nVII, facial nerve; MNs, spinal motor neurons. Scale bar = 100 μm.

upstream of the *Hb9* coding region; in contrast, their corresponding sequences in mouse and human were about 8 kb 5′ of the *Hb9* translation start site (Fig. 4A). This is consistent with *Fugu*'s relatively compact genome. At approximately 400 mb, it is nine times smaller than the mouse or human genome, even though each of these genomes has a similar number of genes (Aparicio et al., 2002; Brenner et al., 1993).

We also analyzed *Hb9* genomic sequence in the zebrafish (*Danio rerio*), and found two copies of the identical region B sequence within 2 kb of the *Hb9* translation start (Fig. 4A). In contrast to *Fugu*, no region A sequence was found in the zebrafish *Hb9* locus, or for that matter anywhere else in the whole zebrafish genome, suggesting that region B might play a more important role than region A in *Hb9* gene expression.

The 125-bp Hb9 enhancer directed motor neuron-selective gene expression in transgenic mice

For functional analysis of the two conserved sequences A and B, we generated several constructs whose relative abilities to drive MN-specific reporter gene expression were assessed in transgenics. Construct #3 included both putative enhancers A and B. It is a 438-bp synthetic chimera that consists of the otherwise noncontiguous 313-bp and 125-bp fragments (Fig. 5A). Transgenic analysis showed that this 438-bp sequence was capable of appropriate temporal and spatial activation of a reporter gene in the ventral spinal cord (Figs. 5B–D). Other than slight *lacZ* expression in the facial nerve, no other ectopic gene expression was noted with this construct at E10.5–11.5. These observations indicate that the *Hb9* regulatory regions contained within the 438-bp chimera of enhancers A and B are sufficient to direct MN-specific gene expression.

Next, we generated constructs #4 and #5, containing three copies of either the 313-bp (fragment A) or 125-bp (fragment B) regions, respectively (Fig. 5A), to investigate whether either of these fragments *alone* was sufficient to direct MN-specific *β-gal* expression. Construct #4 expressed no *β-gal* at E10.5–11.5 (not shown). In contrast, construct #5 expressed *β-gal* in the ventral spinal cord, similar to the expression pattern of constructs #1 and #3 at E10.5–11.5 (Figs. 5E–G). However, ectopic expression of construct #5 was seen in several regions including the ventral midbrain. Together, these results demonstrated that the 125-bp element fragment B was sufficient to direct gene expression to motor neurons, and as such, constituted the *Hb9* core enhancer. However, the 125-bp fragment yielded some ectopic reporter expression as well, and required the additional information supplied by the 313-bp element in order to specify gene expression to MNs.

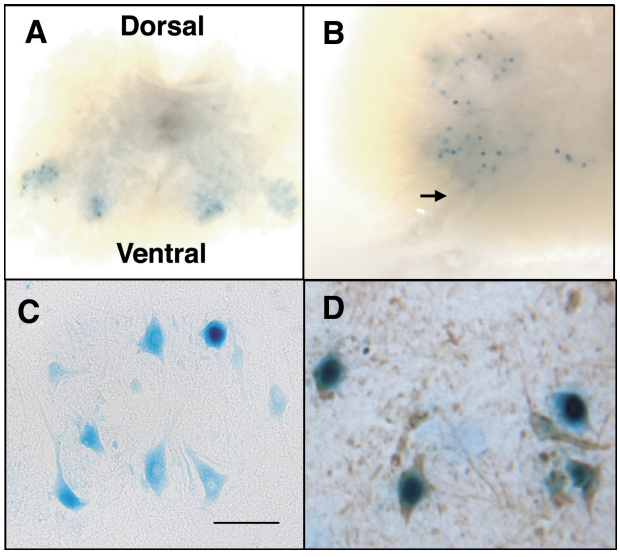
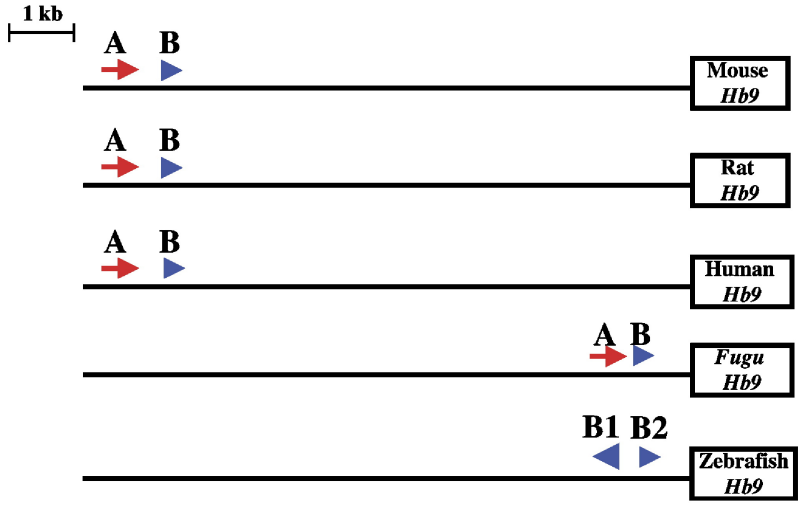


Fig. 3. Expression patterns of reporter genes in the adult spinal cord. (A, B) *β-gal* expressions in representative adult thoracic spinal cord of a 12-week adult transgenic mouse carrying the construct #1. *β-gal* expressions were located in the ventral horn. Arrow indicates extending axons. (C) A transverse section of an adult spinal cord. The *β-gal* activity was restricted to neurons with large cell bodies in the ventral side of spinal, suggesting they are motor neurons. (D) ChAT immunoreactivity (arrowheads) co-localized with *β-gal*.

The requirement of Hox/Pbx sites for the 125-bp Hb9 enhancer activity

To identify potential transcription factor binding sites within the 125-bp *Hb9* enhancer, that might in turn suggest upstream factors important for MN-specific *Hb9* expression, MATInspector (Quandt et al., 1995) was used to search the TRANSFAC database (Heinemeyer et al., 1998). This revealed two sequences (shaded in Fig. 4E), each highly related to a bipartite *Hox*/*Pbx* (HP) consensus binding sequence (TGATNNAT) (Chan and Mann, 1996; Di Rocco et al., 1997) within the 125-bp sequence. These two binding sequences were well conserved among mouse, human, rat, pufferfish, and zebrafish (Fig. 4E). HOX proteins are

A



B

```
MmHb9   AQSNLLGKCRRPRTAFTSQQLLELEHQFKLNKYLSRPKRFEVATSLMLTETQVKIWFQNRRMKWKRSKKAKEQ
RnHb9   ...............................................................................
HsHb9   ...............................................................................
FrHb9   ...............................................................................
DrHb9   ...............................................................................
```

C

	Mouse 313 bp *Hb9* enhancer A	Mouse 125 bp *Hb9* enhancer B
Rat	99%	98%
Human	98%	96%
Fugu	74%	82%
Zebrafish	Not Determined	82%

Fig. 4. Identification of evolutionarily conserved genomic sequences among vertebrate *Hb9* genes. (A) Schematic of *Hb9* genomic region showing the positions and directions of the two highly homologous regions, namely, 313-bp region A (red arrow) and 125-bp region B (blue arrowhead), of the mouse, rat, human, pufferfish, and zebrafish. B1 and B2 indicate two identical copies of the 125-bp sequence found in zebrafish. (B) Comparison of the deduced amino acid sequences of the mouse *Hb9* (*MmHb9*), rat *Hb9* (*RnHb9*), human *Hb9* (*HsHb9*), *Fugu Hb9* (*FrHb9*), zebrafish *Hb9* (*DrHb9*) in the vicinity of the homeodomains. Perfectly conserved homeodomains among all listed species are shaded. (C) The degree of sequence identity of the 313-bp region A and 125-bp region B among evolutionary distinct species. (D, E) Alignment of the 313-bp region A (D) and 125-bp region B (E) from mouse, rat, human, and *Fugu* using the Clustral algorithm. The 125-bp region harbors two *Hox*/*Pbx* consensus binding sequences (TGATNNAT).

D

```
Mouse       TGAATAAATTTAA-GCAGGCTAATTAATATATAAACTAGCTCAATTTGTCAAGTTGATTT
Rat         .............-...........................C...................
Human       .............-GG.........................C...................
Fugu        .TC...CC...TCC.G.A.G..........T.......TTC.G........CT.......

Mouse       GTATTTTAGTTAATTGTGAAAGTAATTACCACATGGTCAAATTAACAGCTTTCTGGAAAT
Rat         ............................................................
Human       ............................................................
Fugu        .........A..........T.............----G.AG.................A.....-

Mouse       GACCAAGCCTGAGGTTTTATTTCCTTCCTGGGTGAAGAAAATTCATTTTTCCAAGCTCTT
Rat         ............................................................
Human       ............................................................
Fugu        ....GG.TTCTG..G.CCGGG...A..TGA..C.CCTT.----.T.A.....TCTC.AA.A

Mouse       GATGTGATGAATAAAAGTCATAAATCTGGGTGATTGGTGCAGGCAGAGTCTAAATGGCTT
Rat         ............................................................
Human       ............................................................
Fugu        .C.........A.GGG.CG..............C....A....CA.TA..........C

Mouse       CATATTTCATTTTAGGTTTAATAGAAATATTCATGCTC--TGTTTTAATGAAATTAAATT
Rat         ...........................................--...............
Human       ...........................................--...............
Fugu        .........CCGCT.C.....................GGGACACC..............CC

Mouse       GAAGGGGGATGGGGCT
Rat         ...............C
Human       ...............C
Fugu        .GC..AA.GC.CA---
```

E

```
Mouse       AGAGTGGTTAGCTGATGAATTGACAAAAACTAATCAGCTTTATTGGGAAACAGGTTTAAG
Rat         .....A.....................................................
Human       ..........................................................-...
Fugu        ..C.C.C....T........C.....................T.G.......-...
Zebrafish   ..................GG..........C..........T.G.......-...

Mouse       GGCACGGACGTGTCAATAACGCTCAGCCTGACCCCCTCTTCCATTAGCT-AGGCAGGCTG
Rat         .......G.............................................-..........
Human       .......G............T.................G..........C..........
Fugu        ....ACTGG..........TT....TTT.....T............A..TTAAGT....T
Zebrafish   ....ACAGA..........TT...TC.......T.T.TC..........T.AAG.....T

Mouse       ATTAGA
Rat         ......
Human       ......
Fugu        ......
Zebrafish   ..A...
```

Fig. 4. (*continued*).

expressed in MNs (Liu et al., 2001; Tiret et al., 1998), and have been implicated as critical determinants of spinal MN identity and organization (Dasen et al., 2003). Thus, it seemed likely that the two conserved bipartite HP motifs within the 125-bp *Hb9* enhancer were involved in activation of the *Hb9* gene. To test this hypothesis, we generated construct #6, by introducing mutations in two HP sites, converting them from TGATNNAT to T**CG**TN**CG**T (Fig. 6A). We confirmed that HOX/PBX proteins failed to form protein–DNA complexes with these mutated HP motifs (Fig. 6E), as previously reported (Maconochie et al., 1997). Mutations of these sequences did not affect cervical *Hb9* gene expression, but completely disrupted more posterior MN-specific gene expression. (Figs. 6B–D). Thus, the HP motifs within the 125-bp *Hb9* enhancer play a critical role in directing MN gene expression in vivo.

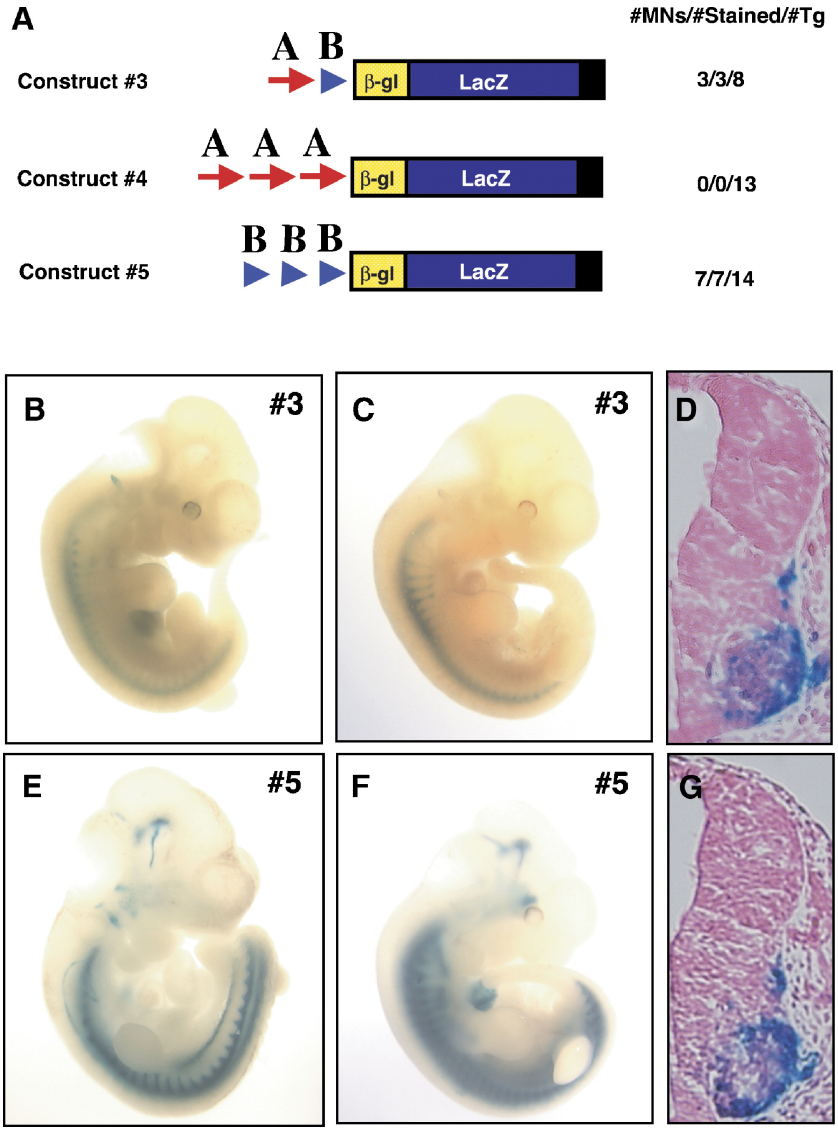


Fig. 5. Functional characterization of the two evolutionarily conserved sequences. (A) Constructs and statistical overview of transgenic embryos. A, 313 bp region A; B, 125 bp region B. (B–E) Expression patterns of reporter genes in representative transgenic embryos. (B–D) The expression of the *β-gal* is detected in spinal MNs of transgenics staged between E11.0 (B) and E11.5 (C, D). (E–G) Staining was observed at the ventral side of spinal cord in the presence of the evolutionarily conserved 125-bp fragments at E10.5 (E) and E11.5 (F, G), whereas the 313-bp fragments did not have any enhancer activity (data not shown).

Cooperative binding of HOXB and PBX1 proteins on HP sites

Next, we tested a representative subset of HOX proteins for cooperative DNA binding with Pbx1, to both the wild-type and a mutated 40-bp sequence derived from the 125-bp *Hb9* enhancer (Fig. 6E). In the presence of PBX1, only HOXB1 and HOXB3 proteins (Fig. 6D, lanes 2 and 3) formed a detectable retarded complex with a DIG-labeled wild-type oligonucleotide. No binding was detected with PBX1 alone (lane 1), or with either of the HOX proteins alone (not shown). Polyclonal antibodies against full-length PBX1 protein completely abolished the formation of the *HoxB1–Pbx1* complex (lane 6). In addition, the introduction of seven point mutations in the core HP motifs abolished cooperative *Hox–Pbx* complex formation (lane 8–11).

Discussion

Evolutionary conservation of Hb9 regulatory elements

Homology searches between genomic sequences of evolutionarily distant species offer a fast detection method for gene regulatory sequences (Muller et al., 2002), and the contracted genomes of the pufferfish and zebrafish have successfully been used to identify conserved regulatory elements in vertebrates (Aparicio et al., 1995). Using this technique, we identified a 3.6-kb fragment first, and then

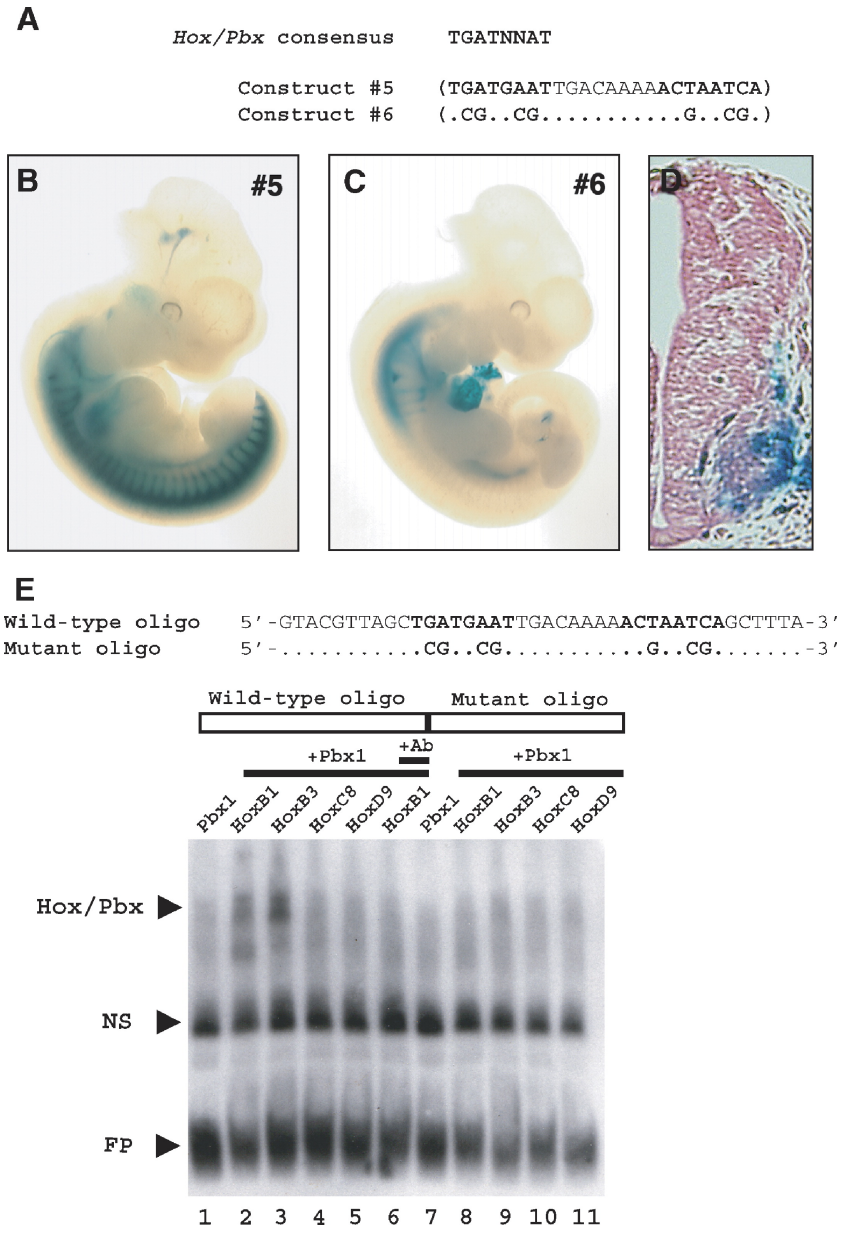


Fig. 6. Identification of a required element for *Hb9* enhancer activity in the spinal cord. (A) Two *Hox*/*Pbx* binding motifs within 125 bp *Hb9* enhancer, and the introduced mutations in construct #6 are indicated. (B) The intact 125 bp *Hb9* enhancer drove *Hb9* expression in cervical, thoracic, and lumbar levels. (C) In vivo inactivation of the 125-bp *Hb9* enhancer by mutations of the two *Hox*/*Pbx* binding motifs. Expression in thoracolumbar MNs was not detected, whereas cervical MN expression was not affected. (D) Low-power coronal section of panel C, taken through cervical cord. (E) EMSA of HOX and PBX1 binding to a 40-bp wild-type or mutant oligonucleotide derived from 125 bp *Hb9* enhancer. Two *Hox*/*Pbx* binding motifs are shown in bold. (Top) Two different oligonucleotides, proteins and antibodies used for binding reactions are indicated. Hox/Pbx indicates the slower migrating complex formed on the wild-type oligonucleotide. FP = free probe. NS = non-specific complexes.

two conserved elements that consisted of noncontiguous 313-bp (region A) and 125-bp (region B) fragments within the 9-kb mouse *Hb9* promoter. Transgenic analysis showed that when these two fragments were serially combined, the resultant 438-bp sequence directed an appropriate temporal and spatial pattern of reporter gene expression in spinal MNs. Further transgenic analysis then showed that the 125-bp region B directed MN-selective expression in vivo, and that the 313 segment was incapable of independently driving reporter gene expression. In this regard, we noted that the zebrafish genome does not contain sequences homologous to the 313-bp region A; rather, BLAST search detected two identical copies of the region B sequence, designated B1 and B2. Together, these data indicate that this 125-bp element represents a motor neuron-specifying *Hb9* core enhancer.

β-gal expression under the control of this enhancer largely coincided with endogenous *Hb9* expression, but there were a few differences. These included a lack of *β-gal* expression in the most caudal segment of spinal cord, and

some ectopic *β-gal* expression in facial nerve, not reflecting endogenous *Hb9* expression (Thaler et al., 1999). Given the difficulty of detecting small fragments of conserved sequences by blast search, especially less than 80 bp, it is possible that there are other functional sequences – both enhancers and silencers – within the 9-kb *Hb9* promoter. In this regard, Lee et al. have recently identified a 246-bp MN enhancer, located from −7135 bp to −6890 bp upstream of the *Hb9* gene (Lee and Pfaff, 2003; Lee et al., 2004), based on its high genomic sequence conservation between human and mouse (>90%). This 246-bp sequence does not overlap with our evolutionarily conserved 125 bp, which is situated −7917 bp to −7793 bp upstream of the *Hb9* coding region. In addition, the 246-bp MN enhancer is not conserved across phylogeny; extensive blast searches failed to identify highly homologous sequences in either the *Fugu* or zebrafish *Hb9* genomic loci. In contrast, the percent identity of our 125-bp *Hb9* enhancer sequence among evolutionarily distinct species, including *Fugu* and zebrafish, exceeded 82%. Together, these studies suggest the existence in mammals of two distinct and non-overlapping MN enhancers, of 125 and 246 bp. Though the 125-bp enhancer is the more phylogenetically conserved, both appear sufficient to direct gene expression to mammalian motor neurons.

Previous studies identified a 4.4-kb *Ngn2* enhancer that exhibited a similar pattern of spinal gene expression as the 125-bp *Hb9* enhancer (Simmons et al., 2001). However, blast search between *Hb9* and *Ngn2* enhancers detected no significant similar sequences, suggesting that the respective expression patterns of these genes are regulated by different upstream transcription factors. Moreover, we conducted blast searches to identify other MN-selective genes that included *Islet1* and *ChAT*, whose regulatory regions might share homology with the *Hb9* 125-bp enhancer. No such blast hit was detected, indicating that the 125-bp enhancer might be regulated uniquely. Unlike the 125-bp region B, the 313-bp region A had no independent enhancer activity. Consistent with this, the zebrafish genome lacks homology to the 313-bp region A. The role of the 313-bp sequence is not clear, but it might function as a silencer that suppresses gene expression in cells other than MNs. Accordingly, we noted that the 125-bp Hb9 enhancer-driven *lacZ* transgenics exhibited increased ectopic *β-gal* expression, compared to the more specific expression directed by the 438-bp chimeric enhancer, which contained the 313-bp region B sequence.

The 3.6-kb Hb9 enhancer directs MN-specific gene expression in the adult spinal cord

Both *Hb9* and *Islet-1* are expressed in adult as well as developing MNs (Vult von Steyern et al., 1999). We thus tested whether our identified *Hb9* enhancer was capable of driving *Hb9* expression in adult MNs. We examined the 3.6-kb *Hb9* enhancer for this purpose, because it confers strong *β-gal* expression in MNs, more so than that afforded by the 438-bp enhancer, and did so without the ectopic expression noted using the 125-bp enhancer alone. In addition, the 3.6-kb segment includes both our 125-bp enhancer and Lee et al.'s 246-bp enhancer. Accordingly, we found that *β-gal* expression in the adult *E/Hb9(3.6kb):lacZ* spinal cord was restricted to magnocellular neurons of the ventral horns, confirming that the 3.6-kb *Hb9* enhancer can target adult MNs. As such, this enhancer may prove useful in establishing MN-targeted vectors, as a means of specifically delivering therapeutic transgenes to vulnerable cells in the motor neuron diseases, such as spinal muscular atrophy and amyotrophic lateral sclerosis. For example, since lentivirus vectors have roughly 8 kb in cloning capacity, the identified 3.6-kb enhancer may be useful to specifically target therapeutic genes of interest (e.g., GDNF, erythropoietin, and/or Bcl-2) to MNs, as a means of selectively supporting those cells and no other. Alternatively, the 3.6-kb *Hb9* enhancer can be used as a MN selection vector, for instance, via fluorescence-activated cell sorting (FACS), using GFP as a reporter (Roy et al., 2004).

A role of Hox/Pbx binding sequences in regulating Hb9 expression

Lee and Pfaff reported that Ngn2/NeuroM transcriptionally synergize with Islet-1 and Lhx3 to directly bind the 250-bp *Hb9* enhancer, thereby activating *Hb9* expression (28). However, most reported regulators of *Hb9* expression, such as *Nkx6.1* and *Olig2*, are transcriptional repressors that do not appear to directly drive *Hb9* expression. Our conserved 125-bp *Hb9* enhancer does not contain E-box sites, suggesting that it may not be regulated by *Ngn2*/*NeuroM*. Rather, we posited that the 125-bp *Hb9* enhancer might direct MN-selective gene expression following binding by other transcription factors, yet-to-be identified.

To better define these alternative molecular pathways of MN induction, we first performed TF binding site analysis of the 125-bp *Hb9* enhancer. This revealed two potential *Hox*/*Pbx* (HP) binding sequences (Fig. 4E) within the 125-bp *Hb9* enhancer, that were also evolutionarily conserved. Along the A–P axis, the nested expression pattern of HOX homeodomain proteins provides positional values that influence the fate of neurons (Lumsden and Krumlauf, 1996), including MNs (Liu et al., 2001; Tiret et al., 1998). Moreover, in the developing hindbrain, misexpression of *Hoxa2* or *Hoxb1* leads to the generation of ectopic MNs (Jungbluth et al., 1999). Our mutagenesis study showed that *Hox*/*Pbx* binding sites were essential for *Hb9* expression in thoracolumbar MNs, suggesting that *Hox*/*Pbx* could be a critical upstream regulator of *Hb9* in these cells. In addition, gel retardation analysis showed that HOXB1 and HOXB3 bound to wild-type sequence derived from the 125-bp *Hb9* enhancer in the presence of PBX1 in vitro, whereas HOXC8 or HOXD9 did not form the retarded complex. These results suggested that *HoxB* genes may play an important role in

regulating *Hb9* expression. Interestingly, *HoxB1*, but not *HoxB3*, *HoxC6*, or *HoxD9*, has previously been shown to bind b1-ARE in vitro, and to regulate *HoxB1* expression (Di Rocco et al., 1997). Yet, HOXB3 bound the 125-bp Hb9 sequence more avidly than did HOXB1. HOXB1 is expressed strongly in the hindbrain where *Hb9* is not expressed, and its spinal cord expression terminates at about E9.5. Thus, HOXB1 might play a role in the initiation of *Hb9* expression, but not in its maintenance after E9.5. On the other hand, HOXB3 appears to be co-expressed with *Hb9* throughout most of its expression domain in the developing spinal cord, though its expression extends more rostrally than that of *Hb9*, suggesting HOXB3 might play the more important role in *Hb9* expression.

Interestingly, TF binding analysis also revealed that one of the two *Hox*/*Pbx* binding sequences contains an overlapping binding sequence of *bicoid*-class homeoprotein (TAATC). However, since none of the known *bicoid*-class homeoproteins including *Otx*, *Goosecoid* (*Gsc*), and *Pitx*, are expressed in the developing spinal cord (Gage et al., 1999; Gaunt et al., 1993; Simeone et al., 1993), it seems unlikely that *bicoid* binding influences MN gene expression, barring the identification of other competent *bicoid*-class homeoproteins with spinal cord expression.

Taken together, these results suggest that the evolutionarily conserved 125-bp region B is a core enhancer that directs gene expression to motor neurons; as such, the molecular pathways regulating *Hb9* expression would appear to be highly conserved across phylogeny. Our data also specifically suggest that HOX/PBX proteins are necessary for the expression of *Hb9* by thoracolumbar MNs, and as such, play a critical role in the segmental specification of spinal MNs. Thus, dorsoventral patterning of the thoracolumbar spinal cord, and Shh-dependent MN specification in particular, would appear to be coordinated with the rostrocaudal patterning of MNs, through HOX/PBX proteins. Besides the developmental insights brought to bear by the identification and analysis of the 125-bp *Hb9* core enhancer, this sequence as well as the 3.6-kb may have great practical and translational value: MN-specifying elements such as the *Hb9* core enhancers may permit us to create not only MN-specific selection vectors, as noted (Roy et al., 2004), but also MN-targeted plasmid and viral vectors for gene therapy. These in turn may permit systemic and whole-CNS administration of gene therapeutics, that may be specifically targeted to spinal MNs using the *Hb9* enhancer.

Acknowledgments

Supported by Project ALS, NINDS, and the Christopher Reeve Paralysis Foundation. We thank Tom Jessell for the *Hb9* promoter and antibody, and Jane Johnson for the BGZA vector. We also thank Drs. Shigeru Kuratani, Takashi Mikawa, and David Reese for their assistance.

References

Aparicio, S., Morrison, A., Gould, A., Gilthorpe, J., Chaudhuri, C., Rigby, P., Krumlauf, R., Brenner, S., 1995. Detecting conserved regulatory elements with the model genome of the Japanese puffer fish, *Fugu rubripes*. Proc. Natl. Acad. Sci. U. S. A. 92, 1684–1688.

Aparicio, S., Chapman, J., Stupka, E., Putnam, N., Chia, J.M., Dehal, P., Christoffels, A., Rash, S., Hoon, S., Smit, A., et al., 2002. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. Science 297, 1301–1310.

Arber, S., Han, B., Mendelsohn, M., Smith, M., Jessell, T.M., Sockanathan, S., 1999. Requirement for the homeobox gene Hb9 in the consolidation of motor neuron identity. Neuron 23, 659–674.

Brenner, S., Elgar, G., Sandford, R., Macrae, A., Venkatesh, B., Aparicio, S., 1993. Characterization of the pufferfish (*Fugu*) genome as a compact model vertebrate genome. Nature 366, 265–268.

Briscoe, J., Ericson, J., 2001. Specification of neuronal fates in the ventral neural tube. Curr. Opin. Neurobiol. 11, 43–49.

Briscoe, J., Sussel, L., Serup, P., Hartigan-O'Connor, D., Jessell, T.M., Rubenstein, J.L., Ericson, J., 1999. Homeobox gene Nkx2.2 and specification of neuronal identity by graded Sonic hedgehog signalling. Nature 398, 622–627.

Briscoe, J., Pierani, A., Jessell, T.M., Ericson, J., 2000. A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell 101, 435–445.

Chan, S.K., Mann, R.S., 1996. A structural model for a homeotic protein–extradenticle–DNA complex accounts for the choice of HOX protein in the heterodimer. Proc. Natl. Acad. Sci. U. S. A. 93, 5223–5228.

Dasen, J.S., Liu, J.P., Jessell, T.M., 2003. Motor neuron columnar fate imposed by sequential phases of Hox-c activity. Nature 425, 926–933.

Di Rocco, G., Mavilio, F., Zappavigna, V., 1997. Functional dissection of a transcriptionally active, target-specific Hox–Pbx complex. EMBO J. 16, 3644–3654.

Gage, P.J., Suh, H., Camper, S.A., 1999. The bicoid-related Pitx gene family in development. Mamm. Genome 10, 197–200.

Gaunt, S.J., Blum, M., De Robertis, E.M., 1993. Expression of the mouse goosecoid gene during mid-embryogenesis may mark mesenchymal cell lineages in the developing head, limbs and body wall. Development 117, 769–778.

Goridis, C., Brunet, J.F., 1999. Transcriptional control of neurotransmitter phenotype. Curr. Opin. Neurobiol. 9, 47–53.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., et al., 1998. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res. 26, 362–367.

Helms, A.W., Abney, A.L., Ben-Arie, N., Zoghbi, H.Y., Johnson, J.E., 2000. Autoregulation and multiple enhancers control Math1 expression in the developing nervous system. Development 127, 1185–1196.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., Pease, L.R., 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51–59.

Hogan, B., Beddington, R., Constantini, F., Lacey, E., 1994. Manipulating the Mouse Embryo: A Laboratory Manual, 2nd edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Jessell, T.M., 2000. Neuronal specification in the spinal cord: inductive signals and transcriptional codes. Nat. Rev., Genet. 1, 20–29.

Jessell, T.M., Melton, D.A., 1992. Diffusible factors in vertebrate embryonic induction. Cell 68, 257–270.

Jungbluth, S., Bell, E., Lumsden, A., 1999. Specification of distinct motor neuron identities by the singular activities of individual Hox genes. Development 126, 2751–2758.

Lee, S.-K., Pfaff, S.L., 2003. Synchronization of neurogenesis and motor neuron specification by direct coupling of bHLH and homeodomain transcription factors. Neuron 38, 731–745.

Lee, S.-K., Jurata, L., Funahashi, J., Ruiz, E., Pfaff, S.L., 2004. Analysis of embryonic motoneuron gene regulation: derepression of general

activators function in concert with enhancer factors. Development 131, 3295–3306.
Liu, J.P., Laufer, E., Jessell, T.M., 2001. Assigning the positional identity of spinal motor neurons: rostrocaudal patterning of Hox-c expression by FGFs, Gdf11, and retinoids. Neuron 32, 997–1012.
Lumsden, A., Krumlauf, R., 1996. Patterning the vertebrate neuraxis. Science 274, 1109–1115.
Maconochie, M.K., Nonchev, S., Studer, M., Chan, S.K., Popperl, H., Sham, M.H., Mann, R.S., Krumlauf, R., 1997. Cross-regulation in the mouse HoxB complex: the expression of Hoxb2 in rhombomere 4 is regulated by Hoxb1. Genes Dev. 11, 1885–1895.
Muhr, J., Andersson, E., Persson, M., Jessell, T.M., Ericson, J., 2001. Groucho-mediated transcriptional repression establishes progenitor cell pattern and neuronal fate in the ventral neural tube. Cell 104, 861–873.
Muller, F., Blader, P., Strahle, U., 2002. Search for enhancers: teleost models in comparative genomic and transgenic analysis of cis regulatory elements. Bioessays 24, 564–572.
Ovitt, C.E., Yoem, Y.I., Scholer, H.R., 1997. Transgenic analysis of embryonic gene expression using LacZ as a reporter. Microinjection and Transgenesis. 427–437.
Pierani, A., Moran-Rivard, L., Sunshine, M.J., Littman, D.R., Goulding, M., Jessell, T.M., 2001. Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. Neuron 29, 367–384.
Quandt, K., Frech, K., Karas, H., Wingender, E., Werner, T., 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23, 4878–4884.
Roy, N., Nakano, T., Keyoung, H., Windrem, M., Rashbaum, W., Alonso, M., Kang, J., Peng, W., Carpenter, M., Lin, J., et al., 2004. Telomerase-immortalization of neuronally restricted progenitor cells derived from the human fetal spinal cord. Nat. Biotechnol. 22, 297–305.
Simeone, A., Acampora, D., Mallamaci, A., Stornaiuolo, A., D'Apice, M.R., Nigro, V., Boncinelli, E., 1993. A vertebrate gene related to orthodenticle contains a homeodomain of the bicoid class and demarcates anterior neuroectoderm in the gastrulating mouse embryo. EMBO J. 12, 2735–2747.
Simmons, A.D., Horton, S., Abney, A.L., Johnson, J.E., 2001. Neurogenin2 expression in ventral and dorsal spinal neural tube progenitor cells is regulated by distinct enhancers. Dev. Biol. 229, 327–339.
Tanabe, Y., William, C., Jessell, T.M., 1998. Specification of motor neuron identity by the MNR2 homeodomain protein. Cell 95, 67–80.
Thaler, J., Harrison, K., Sharma, K., Lettieri, K., Kehrl, J., Pfaff, S.L., 1999. Active suppression of interneuron programs within developing motor neurons revealed by analysis of homeodomain factor HB9. Neuron 23, 675–687.
Tiret, L., Le Mouellic, H., Maury, M., Brulet, P., 1998. Increased apoptosis of motoneurons and altered somatotopic maps in the brachial spinal cord of Hoxc-8-deficient mice. Development 125, 279–291.
Vallstedt, A., Muhr, J., Pattyn, A., Pierani, A., Mendelsohn, M., Sander, M., Jessell, T.M., Ericson, J., 2001. Different levels of repressor activity assign redundant and specific roles to Nkx6 genes in motor neuron and interneuron specification. Neuron 31, 743–755.
Vult von Steyern, F., Martinov, V., Rabben, I., Nja, A., de Lapeyriere, O., Lomo, T., 1999. The homeodomain transcription factors Islet 1 and HB9 are expressed in adult alpha and gamma motoneurons identified by selective retrograde tracing. Eur. J. Neurosci. 11, 2093–2102.
Wichterle, H., Lieberam, I., Porter, J.A., Jessell, T.M., 2002. Directed differentiation of embryonic stem cells into motor neurons. Cell 110, 385–397.
Yee, S.P., Rigby, P.W., 1993. The regulation of myogenin gene expression during the embryonic development of the mouse. Genes Dev. 7, 1277–1289.